

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 06:56:21 ; Search time 149,542 Seconds
(without alignments)
1349.183 Million cell updates/sec

Title: US-09-939-537-29_COPY_1_394
Perfect score: 2029
Sequence: 1 NMRGVFRRHLVLQLALLP.....SGVLLSNIKVLPTWSTPV 394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2015	99.3	458	1	CD4_HUMAN
2	1991	98.1	458	1	CD4_PANTR
3	1844	90.9	458	1	CD4_MACFU
4	1843	90.8	458	1	CD4_MACMU
5	1840	90.7	458	1	CD4_MACFA
6	1835	90.4	458	1	CD4_MACNE
7	1818	89.6	458	1	CD4_CERAE
8	1717	84.6	397	1	CD4_BRYPA
9	1707	84.1	397	1	CD4_CERTO
10	1705	84.0	397	2	009261
11	1701	83.8	397	2	009260
12	1700	83.8	397	2	009259
13	1696	83.6	397	2	009259
14	1696	83.6	397	2	009263
15	1689	83.2	397	2	009262
16	1601.5	78.9	457	2	08H2T8
17	1575.5	77.6	457	2	08H2T7
18	1572.5	77.5	457	2	08H2T7
19	1183.5	58.3	455	1	CD4_SAISC
20	1182.5	58.3	455	2	0710E2
21	1146	56.5	459	1	CD4_RABIT
22	1135	55.9	463	1	CD4_CANFA
23	1119.5	55.2	457	2	06G1R3
24	1099	54.2	432	2	06LBN1
25	1097.5	54.1	444	2	06LBN4
26	1087	53.6	474	2	P79355
27	1078.5	53.2	406	1	CD4_RAT
28	999	49.2	457	1	CD4_MOUSE
29	993	48.9	457	1	CD4_MOUSE
30	985	48.5	457	2	061396
31	954.5	47.0	433	2	055054

32	400	19.7	86	2	077596
33	400	19.7	86	2	077597
34	398	19.6	78	2	06LCP8
35	397	19.6	86	2	077594
36	397	19.6	86	2	077599
37	395	19.5	86	2	077595
38	395	19.5	86	2	0710S5
39	392	19.3	86	2	077598
40	388	19.1	86	2	077601
41	383	18.9	86	2	077600
42	357	17.6	71	2	013969
43	305.5	15.1	99	2	029027
44	280.5	13.8	99	2	029028
45	271.5	13.4	482	2	090WB5

ALIGNMENTS

RESULT 1
CD4_HUMAN STANDARD; PRT; 458 AA.
ID CD4_HUMAN
AC P01730;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
GN Name=CD4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85254948; PubMed=2990730;
RA Maddon P.J., Littman D.R., Godfrey M., Maddon D.E., Chess L., Axel R.;
RT "The isolation and nucleotide sequence of a cDNA encoding the T cell
RT surface protein T4: a new member of the immunoglobulin gene family.";
RL Cell 42:93-104(1985).
RN [2]
RP REVISION TO 26.
RX MEDLINE=89028665; PubMed=3263213; DOI=10.1016/0092-8674(88)90211-5;
RA Littman D.R., Maddon P.J., Axel R.;
RT "Corrected CD4 sequence.";
RL Cell 55:541-541(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96303695; PubMed=8723724;
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
RA Mailey T., Gibbs R.A.;
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
RT genes at human chromosome 12p13.";
RL Genome Res. 6:314-326(1996).
RN [4]
RP SEQUENCE FROM N.A. AND VARIANT TRP-265.
RX MEDLINE=91216786; PubMed=1708753; DOI=10.1016/0198-8859(91)90077-M;
RA Hodge T.W., Sasso D.R., McDougal J.S.;
RT "Humans with OKT4-epitope deficiency have a single nucleotide base
RT change in the CD4 gene, resulting in substitution of TRP240 for
RT ARG240.";
RL Hum. Immunol. 30:99-104(1991).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Scrusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL [6]
 RN SEQUENCE OF 28-424 FROM N.A.
 RP TISSUE=Blood;
 RX MEDLINE=93049640; PubMed=1425921;
 RT Fomsgard A., Hirsch V.M., Johnson P.R.,
 RT "Cloning and sequences of primate CD4 molecules: diversity of the
 RT cellular receptor for simian immunodeficiency virus/human
 RT immunodeficiency virus.";
 RL Eur. J. Immunol. 22:2973-2981 (1992).
 RL [7]
 RN SEQUENCE OF 26-394, AND DISULFIDE BOND.
 RP MEDLINE=90078232; PubMed=2592374;
 RX Carr S.A., Hemling M.E., Folera-Wasserman G., Sweet R.W., Anumula K.,
 RA Baur J.R., Huddleston M.J., Taylor P.,
 RA "Protein and carbohydrate structural analysis of a recombinant soluble
 RT CD4 receptor by mass spectrometry.";
 RT J. Biol. Chem. 264:21286-21295 (1989).
 RN [8]
 RN SEQUENCE OF 26-40.
 RX PubMed=15340161; DOI=10.1110/ps.04682504;
 RA Zhang Z., Henzel W.U.,
 RA "Signal peptide prediction based on analysis of experimentally
 RT verified cleavage sites.";
 RT Protein Sci. 13:2819-2824 (2004).
 RN [9]
 RN REMOVAL FROM CELL SURFACE BY HIV-1 NEF, AND MUTAGENESIS OF MET-432;
 RP SER-433; 438-LEU-439 AND SER-440.
 RX PubMed=8124721;
 RA Aiken C., Konner J., Landau N.R., Lenburg M.E., Trono D.,
 RA "Nef induces CD4 endocytosis: requirement for a critical dileucine
 RT motif in the membrane-proximal CD4 cytoplasmic domain.";
 RT Cell 76:853-864 (1994).
 RN [10]
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 26-208.
 RP MEDLINE=91061881; PubMed=1701030; DOI=10.1038/348411a0;
 RX Wang J., Yan Y., Garrett T.P., Liu J., Rodgers D.W., Garlick R.L.,
 RA Tarr G.E., Husain Y., Reinherz E.L., Harrison S.C.,
 RA "Atomic structure of a fragment of human CD4 containing two
 RT immunoglobulin-like domains.";
 RT Nature 348:411-418 (1990).
 RN [11]
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-208.
 RP MEDLINE=91061882; PubMed=2247146; DOI=10.1038/348419a0;
 RX Ryu S.-E., Kwong P.D., Truneh A., Porter T.G., Arthos J.,
 RA Rosenberg M., Dai X., Xuong N.-H., Axel R., Sweet R.W.,
 RA Hendrickson W.A.,
 RA "Crystal structure of an HIV-binding recombinant fragment of human
 RT CD4.";
 RL Nature 348:419-426 (1990).
 RN [12]
 RN X-RAY CRYSTALLOGRAPHY (3.9 ANGSTROMS) OF 26-388.
 RP MEDLINE=97311402; PubMed=9168119;
 RX Wu H., Kwong P.D., Hendrickson W.A.,
 RT "Dimeric association and segmental variability in the structure of
 RT human CD4.";
 RL Nature 387:527-530 (1997).
 RN [13]
 RN PALMITOYLATION.
 RP MEDLINE=92317088; PubMed=1618861;
 RX Crise B., Rose J.K.,

RT "Identification of palmitoylation sites on CD4, the human
 RT immunodeficiency virus receptor.";
 RL J. Biol. Chem. 267:13593-13597 (1992).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Removed from plasma
 CC membrane by HIV-1 Nef protein that increases clathrin-dependent
 CC endocytosis of this antigen to target it to lysosomal degradation.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like V-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD4 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd4.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ldb-rib.ch/announce/>
 CC or send an email to license@ldb-rib.ch).
 CC -----
 CC EMBL; M12807; AAA35572.1; -;
 DR EMBL; U47924; AAB51309.1; -;
 DR EMBL; M35160; AAB16069.1; -;
 DR EMBL; BC025782; AAB25782.1; -;
 DR PIR; A90872; RWHUT4.
 DR PDB; 1CDH; X-ray; @=26-203.
 DR PDB; 1CDI; X-ray; @=25-203.
 DR PDB; 1CDJ; X-ray; @=26-203.
 DR PDB; 1CDU; X-ray; @=26-203.
 DR PDB; 1CDY; X-ray; @=26-203.
 DR PDB; 1G9M; X-ray; C=26-210.
 DR PDB; 1G9N; X-ray; C=26-210.
 DR PDB; 1GCI; X-ray; C=26-210.
 DR PDB; 1TL4; X-ray; D=26-203.
 DR PDB; 1O68; NMR; A=421-458.
 DR PDB; 1WBR; NMR; @=427-445.
 DR PDB; 1WIO; X-ray; A/B=26-388.
 DR PDB; 1WIP; X-ray; A/B=26-388.
 DR PDB; 1WIQ; X-ray; A/B=26-388.
 DR PDB; 3CD4; X-ray; @=26-207.
 DR GlycoSuiteDB; P01730; -;
 DR GeneW; HGNC:1678; CD4.
 DR H-IvDB; HIX0023001; -;
 DR MIM; 186940; -;
 DR GO; GO:0005886; C:Plasma membrane; TAS.
 DR GO; GO:0042101; C:T-cell receptor complex; NAS.
 DR GO; GO:0015026; F:coreceptor activity; NAS.
 DR GO; GO:0015029; F:internalization receptor activity; TAS.
 DR GO; GO:0042289; F:MHC class II protein binding; NAS.
 DR GO; GO:0004888; F:transmembrane receptor activity; TAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; NAS.
 DR GO; GO:0030217; P:T-cell differentiation; NAS.
 DR GO; GO:0045058; P:T-cell selection; NAS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; NAS.
 DR InterPro; IPR008424; CD2.
 DR InterPro; IPR000973; CD4 TCAG.
 DR InterPro; IPR007110; Ig-Like.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_2.
 DR PRINTS; PR00692; CD4CANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 KW 3D-structure; Direct protein sequencing; Glycoprotein;
 KW Immune response; Immunoglobulin domain; Lipoprotein; Palmitate;
 KW Polymorphism; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 25
 FT CHAIN 26 458 T-cell surface glycoprotein CD4.
 FT DOMAIN 26 396 Extracellular (Potential).
 FT TRANSMEM 397 418 Potential.
 FT DOMAIN 419 458 Cytoplasmic (Potential).

```

FT DOMAIN 26 125 Ig-like V-type.
FT DOMAIN 126 203 Ig-like C2-type 1.

Query Match 99.3%; Score 2015; DB 1; Length 458;
Best Local Similarity 99.5%; Pred. No. 2,8e-139;
Matches 392; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NMRGVPFRLHLVLTALPAATQGNKVVLLGKGGDTVELTCTASQKSIQFHWKSNQIK 60
DB 1 NMRGVPFRLHLVLTALPAATQGNKVVLLGKGGDTVELTCTASQKSIQFHWKSNQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNPFLIIKNIKIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNPFLIIKNIKIEDSDTYICEVEDQKEEVL 120
QY 121 LVFGITANSPTHTLLOGOSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGITANSPTHTLLOGOSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNKQKVEPKIDIVLAFQKASIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEPKIDIVLAFQKASIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKCLPLHLTLPQALPOVAGSGNLTIA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKCLPLHLTLPQALPOVAGSGNLTIA 300
QY 301 LEAKTGKHLQHEVNLVYMRATQLOKNTLCEVNGPTSPKMLSLKLENKAKVSKREKPVWV 360
DB 301 LEAKTGKHLQHEVNLVYMRATQLOKNTLCEVNGPTSPKMLSLKLENKAKVSKREKPVWV 360
QY 361 LNPEAGMMOCLLSDSGQVLLSESNIKVLPWSTPVP 394
DB 361 LNPEAGMMOCLLSDSGQVLLSESNIKVLPWSTPVP 394

RESULT 2
CD4_PANTR STANDARD; PRT; 458 AA.
AC P16004;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
GN Name=CD4;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90182664; PubMed=2107024; DOI=10.1016/0092-8674(90)90089-W;
RA Camerini D., Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE OF 26-424 FROM N.A.
RX TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL, M3135; AAA35407.1; -
DR EMBL, X73323; CAA51749.1; -.
DR F01, B37322; RMC274.
DR HSSP; P01730; 1W10.
DR GO; GO:0042101; C: T-cell receptor complex; ISS.
DR GO; GO:0015026; F: coreceptor activity; ISS.
DR GO; GO:0042289; F: MHC class II protein binding; ISS.
DR GO; GO:0006955; P: immune response; ISS.
DR GO; GO:0045086; P: positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P: T-cell differentiation; ISS.
DR GO; GO:0007169; P: transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR000973; CD4 TCAG.
DR InterPro; IPR007110; Ig_V.
DR InterPro; IPR003596; Ig_L.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_Like; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 126 125
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT CONFLICT 42 62
FT CONFLICT 62 62
FT CONFLICT 191 191
SQ SEQUENCE 458 AA; 51057 MW; A7C3ACBA5257D3AD CRC64;

Query Match 98.1%; Score 1991; DB 1; Length 458;
Best Local Similarity 98.5%; Pred. No. 1,6e-137;
Matches 388; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 NMRGVPFRLHLVLTALPAATQGNKVVLLGKGGDTVELTCTASQKSIQFHWKSNQIK 60
DB 1 NMRGVPFRLHLVLTALPAATQGNKVVLLGKGGDTVELTCTASQKSIQFHWKSNQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNPFLIIKNIKIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNPFLIIKNIKIEDSDTYICEVEDQKEEVL 120
QY 121 LVFGITANSPTHTLLOGOSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGITANSPTHTLLOGOSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNKQKVEPKIDIVLAFQKASIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEPKIDIVLAFQKASIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKCLPLHLTLPQALPOVAGSGNLTIA 300

```

Db 241 QAERASSKSWITFDLKNKEVSVKRVTDPKLQMGKLLPLHLTPOLPOYAGSGNLTLLA 300
 QY 301 LEAKTGKHQEVNLYVWRATQKLTCEVWGPTSPKMLSLKLENKAKVSKREKVVW 360
 Db 301 LEAKTGKHQEVNLYVWRATQKLTCEVWGPTSPKMLSLKLENKAKVSKREKVVW 360
 QY 361 INPEAGMWCILSDSGQVLTLESNIKVLPTWSTPV 394
 Db 361 INPEAGMWCILSDSGQVLTLESNIKVLPTWSTPV 394

RESULT 3
 CD4_MACFU STANDARD: PRT; 458 AA.

AC P79184;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
 GN Name=CD4;
 OS Macaca fuscata fuscata (Japanese macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
 OC NCBI_Taxid=9543;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hashimoto O., Tatum M.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.
 CC - SUBUNIT: Associates with p56-ck (By similarity).
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; D63348; BA09672.1; -
 DR HSBP; P01730; IMBR.
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro; IPR008424; CD2.
 DR InterPro; IPR000973; CD4_TCAg.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein; Palmitate; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 25 By similarity.
 FT CHAIN 26 458 T-cell surface glycoprotein CD4.
 FT DOMAIN 26 396 Extracellular (Potential).
 FT TRANSMEM 397 418 Potential.
 FT DOMAIN 419 458 Cytoplasmic (Potential).
 FT DOMAIN 26 125 Ig-like V-type.
 FT DOMAIN 126 203 Ig-like C2-type 1.
 FT DOMAIN 204 317 Ig-like C2-type 2.

FT DOMAIN 318 374 Ig-like C2-type 3.
 FT CARBOHYD 42 42 N-linked (GlcNAc...) (potential).
 FT CARBOHYD 296 296 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 325 325 N-linked (GlcNAc...) (By similarity).
 FT DISULFID 41 109 By similarity.
 FT DISULFID 155 184 By similarity.
 FT DISULFID 328 370 By similarity.
 FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
 FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
 FT SEQUENCE 458 AA; 50828 MW; 76B3E7EF08185535 CRC64;

Query Match 90.9%; Score 1844; DB 1; Length 458;
 Best Local Similarity 90.4%; Pred. No. 9,5e-127; Indels 0; Gaps 0;
 Matches 356; Conservative 17; Mismatches 21;

QY 1 MNRGVPFRHLLVLTQALLPAAATQGNKVVLTGKGDIVETCTASQKKSIOFHKNSNQIK 60
 Db 1 MNRGIPFRHLLVLTQALLPAAATQGNKVVLTGKGDIVETCTASQKKSIOFHKNSNQIK 60
 QY 61 ILGNQGSFLTKGPSKINDRADSRSLMDQGNPPLIKNLKIEDSDTYICBVDQKEVQL 120
 Db 61 ILGNQGSFLTKGPSKINDRADSRSLMDQGNPPLIKNLKIEDSDTYICBVDQKEVQL 120
 QY 121 LVFGLTANSSTHLLQGSLLTLTSPGPSVQCSPPGKNIQGGKTLVSQLEQDSG 180
 Db 121 LVFGLTANSSTHLLQGSLLTLTSPGPSVQCSPPGKNIQGGKTLVSQLEQDSG 180
 QY 181 TWTCVTQONQKVEFKIDIVLAFQKASIVYKKEGQVFEFSPPLAFTVEKLTGSGELW 240
 Db 181 TWTCVTQONQKVEFKIDIVLAFQKASIVYKKEGQVFEFSPPLAFTVEKLTGSGELW 240
 QY 241 QAERASSKSWITFDLKNKEVSVKRVTDPKLQMGKLLPLHLTPOLPOYAGSGNLTLLA 300
 Db 241 QAERASSKSWITFDLKNKEVSVKRVTDPKLQMGKLLPLHLTPOLPOYAGSGNLTLLA 300
 QY 301 LEAKTGKHQEVNLYVWRATQKLTCEVWGPTSPKMLSLKLENKAKVSKREKVVW 360
 Db 301 LEAKTGKHQEVNLYVWRATQKLTCEVWGPTSPKMLSLKLENKAKVSKREKVVW 360
 QY 361 INPEAGMWCILSDSGQVLTLESNIKVLPTWSTPV 394
 Db 361 INPEAGMWCILSDSGQVLTLESNIKVLPTWSTPV 394

RESULT 4
 CD4_MACFU STANDARD: PRT; 458 AA.

AC P16003; Q29617;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
 GN Name=CD4;
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
 OC NCBI_Taxid=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90182664; PubMed=2107024; DOI=10.1016/0092-8674(90)90089-W; Camerini D., Seed B.;
 RT "A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site";
 RT Cell 60:747-754(1990).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Thymocytes;
 RA Hashimoto O., Tatum M.;
 RT "Molecular cloning and expression of macaque CD4s";
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE OF 28-424 FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=93049640; PubMed=1425921;
 RA Fomgaard A., Hirsch V.M., Johnson P.R.;
 RT "Cloning and sequences of primate CD4 molecules: diversity of the
 RT cellular receptor for simian immunodeficiency virus/human
 RT immunodeficiency virus.";
 RL Eur. J. Immunol. 22:2973-2981(1992).
 RN [4]
 RP SEQUENCE OF 107-192 FROM N.A.
 RX MEDLINE=98320644; PubMed=9656488;
 RA Harris B.E., Disocell T.R.;
 RT "Nuclear gene trees and the phylogenetic relationships of the
 RT mangabeys (Primates: Papionini).";
 RL Mol. Biol. Evol. 15:692-900(1998).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, M3134; AAA6838.1; -;
 DR EMBL, D63347; BAA09671.1; -;
 DR EMBL, X73326; CAA51752.1; -;
 DR EMBL, AF057385; AAC25129.1; -;
 DR HSP, P01730; 1MR.
 DR GO, GO:0042101; C:T-cell receptor complex; ISS.
 DR GO, GO:0015026; P:coreceptor activity; ISS.
 DR GO, GO:0042289; F:MHC class II protein binding; ISS.
 DR GO, GO:0005955; P:immune response; ISS.
 DR GO, GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
 DR GO, GO:0030217; P:T-cell differentiation; ISS.
 DR GO, GO:0045058; P:T-cell selection; ISS.
 DR GO, GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
 DR InterPro: IPR008434; CD2.
 DR InterPro: IPR000973; CD4_TCRg.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam; PR00047; Ig_2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 25
 FT CHAIN 26 458
 FT DOMAIN 26 396
 FT TRANSMEM 397 418
 FT DOMAIN 419 458
 FT DOMAIN 458 458
 FT DOMAIN 126 203
 FT DOMAIN 204 317
 FT DOMAIN 318 374
 FT CARBOHYD 296 296
 FT CARBOHYD 325 325
 FT DISULFID 41 109
 FT DISULFID 155 155
 FT DISULFID 328 370
 FT LIPID 419 419
 FT LIPID 422 422
 FT CONFLICT 42 42
 FT CONFLICT 62 62
 FT CONFLICT 67 67
 FT CONFLICT 169 169
 I -> L (in Ref. 2).

FT CONFLICT 191 191 K -> N (in Ref. 3).
 FT CONFLICT 248 248 S -> P (in Ref. 2).
 FT CONFLICT 265 265 R -> Q (in Ref. 3).
 FT CONFLICT 349 349 A -> T (in Ref. 2).
 SQ SEQUENCE 458 AA; 50884 MW; 8BB80339FAFEC808 CRC64;
 Query Match 90.8%; Score 1843; DB 1; Length 458;
 Best Local Similarity 90.4%; Pred. No. 1,le-126;
 Matches 356; Conservative 17; Mismatches 21; Indels 0; Gaps 0;
 QY 1 MNRGVPFRLHLLVQLALPAATQGNKVVLAGKGDVETLTASQKKSIOFHWNKNQIK 60
 DB 1 MNRGIPFRLHLLVQLALPAATQGNKVVLAGKGDVETLTASQKKSIOFHWNKNQIK 60
 QY 61 IIGNGSFLTKGPKSLNDPADSRSLMDGNSPPLIKIKIDSPYITGVDDQKEVQL 120
 DB 61 IIGIQGLFTTKPKSLSDPADSRSLMDGNSPPLIKIKIDSPYITGVDDQKEVQL 120
 QY 121 LVFGLTANSDTHLQGSQTLTLFESPGSSPVSQCRSPGKNIOGAGTSLVQLQLQDSG 180
 DB 121 LVFGLTANSDTHLQGSQTLTLFESPGSSPVSQCRSPGKNIOGAGTSLVQLQLQDSG 180
 QY 181 TWCTVQLQNKVKEFKIDIVLAFOKASSIVYKKEGQVFSPLAFTVEXLTGSGELMW 240
 DB 181 TWCTVSGDQKTVKEFKIDIVLAFOKASSIVYKKEGQVFSPLAFTVEXLTGSGELMW 240
 QY 241 QABRASSKSWITFDPLKKEVSVKRVTQDPKQNGKKPLHLTLPOALPOVAGSGNLTLA 300
 DB 241 QABRASSKSWITFDPLKKEVSVKRVTQDPKQNGKKPLHLTLPOALPOVAGSGNLTLA 300
 QY 301 LEAKTGKHOENVLVVWRATQLOKNUJCEWGPSTPKMLSLKLEKAEVSKREKPVWV 360
 DB 301 LEAKTGKHOENVLVVWRATQLOKNUJCEWGPSTPKMLSLKLEKAEVSKREKPVWV 360
 QY 361 LNPEAGMQLSDSGQVLLESINIKVPTWPTPV 394
 DB 361 LNPEAGMQLSDSGQVLLESINIKVPTWPTPV 394
 RESULT 5
 ID CD4_MACPA STANDARD; PRT; 458 AA.
 AC P79185;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 05-JUL-1998 (Rel. 36, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
 DE T4/Len-3).
 GN Name=CD4;
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymocytes;
 RA Tatsumi M., Yabe M., Yamada Y.K.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC -----
DR EMBL: D63349; BAA09673.1; -.
DR HSSP: P01730; IMBR.
DR GO: GO:0042101; C-T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4 TCAG.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
SQ SEQUENCE 458 AA; 50872 MW; 9105479F85C56FF7 CRC64;

Query Match 90.7%; Score 1840; DB 1; Length 458;
Best Local Similarity 90.4%; Pred. No. 1.9e-126;
Matches 356; Conservative 16; Mismatches 22; Indels 0; Gaps 0;

1 NMRGPVFRHLVLTOLALPAATQGNKVKYLGKKGDTVELTCTASOKKSIQPHKNSNOIK 60
1 NMRGIPFRHLVLTOLALPAATQGNKVKYLGKKGDTVELTCTASOKKSIQPHKNSNOIK 60
61 ILGNQGSFLTQPSKLNDRADRSRSLMDQGNFPLIKNLKIDSDTYICEVEDEKERVQL 120
61 ILGIQGSFLTQPSKLNDRADRSRSLMDQGCFSMIIKNLKIDSDTYICEVENKKEVEL 120
121 LVFGILTANDTHLLQGOSITLTLESPPGSSPSVQCRSPGKNIOGAKTYSVQLEIDSG 180
121 LVFGILTANDTHLLHGQSITLTLESPPGSSPSVKCSPGGKNIOGAKTYSVQLEIDSG 180
121 LVFGILTANDTHLLHGQSITLTLESPPGSSPSVKCSPGGKNIOGAKTYSVQLEIDSG 180
181 TWTCTVLQNGQKVEFPKIDIVLAFQKASIVYKKSGEYVFSPLAFTTEKLTGSGELMW 240
181 TWTCTVSDQKVEFPKIDIVLAFQKASIVYKKSGEYVFSPLAFTTEKLTGSGELMW 240
241 QAERASSSKSWITFPDKNKEVSQKRVTDPKLQMGKKLPLHLTLPOLAPQVAGSGLTLA 300
241 QAERASSSKSWITFPDKNKEVSQKRVTDPKLQMGKKLPLHLTLPOLAPQVAGSGLTLA 300
301 LEAKTGKGLHGVNLVYMRATQLOKNTLGVWGPSTSPKMLSLTKENKAQVSKREKPVVW 360
301 LEAKTGKGLHGVNLVYMRATQLOKNTLGVWGPSTSPKMLSLTKENKAQVSKREKPVVW 360
361 INPEAGMOCILSDSGQVLLESNIKVLPTWSTPV 394
361 INPEAGMOCILSDSGQVLLESNIKVLPTWSTPV 394

```

```

RESULT 6
CD4_MACNE STANDARD; PRT; 458 AA.
ID CD4_MACNE
AC 008340; P79196;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-FEB-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN Name=CD4;
OS Macaca nemestrina (pig-tailed macaque).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eumartalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_Taxid=9545;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RL submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC - SUBUNIT: Associates with p56-lck (By similarity).
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D63346; BAA09670.1; -.
DR EMBL: X73325; CAA51751.1; -.
DR HSSP: P01730; IMBR.
DR GO: GO:0042101; C-T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4 TCAG.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
SQ SEQUENCE 458 AA; 50872 MW; 9105479F85C56FF7 CRC64;

```

```

FT CARBOHYD 42 42 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 296 296 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc...) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 328 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 57 57 D -> N (in Ref. 2).
FT CONFLICT 91 91 C -> H (in Ref. 2).
FT CONFLICT 105 105 N -> D (in Ref. 2).
FT CONFLICT 113 113 N -> B (in Ref. 2).
FT CONFLICT 302 302 D -> B (in Ref. 2).
FT CONFLICT 349 349 T -> A (in Ref. 2).
SQ SEQUENCE 458 AA; 50905 MW; 751A9BA2C8B3EB16 CRC64;

Query Match 90.4%; Score 1835; DB 1; Length 458;
Best Local Similarity 89.8%; Pred. No. 4.3e-126;
Matches 354; Conservative 19; Mismatches 21; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLVLTQALPAATQGNKVVLAGKGDVVELTCTASOKKSTQFMKNSNOIK 60
DB 1 MNRGVPFRLHLVLTQALPAATQGNKVVLAGKGDVVELTCTASOKKSTQFMKNSNOIK 60
QY 61 ILAGNOSFLLTKGPSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEVQI 120
DB 61 ILAGNOSFLLTKGPSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEVQI 120
QY 121 LVFGLTANSDTHLQOSITLTLESPPGSSPSVQCSPPGKNIQSGKTLTSSVQLQDSG 180
DB 121 LVFGLTANSDTHLQOSITLTLESPPGSSPSVQCSPPGKNIQSGKTLTSSVQLQDSG 180
QY 181 TWTCTVLOKQKVEFKIDIVLAFOKASSIVYKKEBOVBPSPPLAFYTEKLTGSGELWM 240
DB 181 TWTCTVLOKQKVEFKIDIVLAFOKASSIVYKKEBOVBPSPPLAFYTEKLTGSGELWM 240
QY 241 QAERASSSSKWTTFDKNKEVSKRTQPKLOMGKKLPPLHLTPQALPOYAGSGNLTTA 300
DB 241 QAERASSSSKWTTFDKNKEVSKRTQPKLOMGKKLPPLHLTPQALPOYAGSGNLTTA 300
QY 301 LEAKTGKHOEVNLVVMRAITQLOKNTLCEVWGFSTPKMLSLKLNKKAQVSRKRPVV 360
DB 301 LEAKTGKHOEVNLVVMRAITQLOKNTLCEVWGFSTPKMLSLKLNKKAQVSRKRPVV 360
QY 361 LNPEAGMOCILSDSGVLLSENIKTLPTWSTPV 394
DB 361 LNPEAGMOCILSDSGVLLSENIKTLPTWSTPV 394

RESULT 7
CD4_CERAE STANDARD; PRT; 458 AA.
ID CD4_CERAE STAN 458
DC 01-FEB-1995 (Rel. 31, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
GN Name=CD4;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OC NCBI_Taxid=9534;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatum M.;
RT "Molecular cloning and expression of african green monkey CD4.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;

```

```

RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981 (1992).
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Peripheral blood;
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Machlot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128 (1997).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=956488;
RA Harris E.E., Disocell T.R.;
RT "Nuclear gene trees and the phylogenetic relationships of the
RT mangabeys (Primates: Papionini).";
RL Mol. Biol. Evol. 15:892-900 (1998).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86589; BAA1332.1; -
DR EMBL; X73322; CAA51748.1; -
DR EMBL; AF001226; AAB60875.1; -
DR EMBL; AF001228; AAB60875.1; -
DR EMBL; AF057380; AAC25124.1; -
DR HSSP; P01730; 1WIO.
DR GO; GO:0042101; C: T-cell receptor complex; ISS.
DR GO; GO:0015026; F: coreceptor activity; ISS.
DR GO; GO:0043289; F: MHC class II protein binding; ISS.
DR GO; GO:0006955; P: immune response; ISS.
DR GO; GO:0045086; P: positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P: T-cell differentiation; ISS.
DR GO; GO:0045058; P: T-cell selection; ISS.
DR GO; GO:0007169; P: transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR009773; CD4_TCRG.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT TRANSMEM 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 281 281 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 296 296 N-linked (GlcNAc...) (Potential).
FT DISULFID 325 325 N-linked (GlcNAc...) (Potential).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.

```

FT DISULFID 328 370 By similarity.
 FT LIPID 419 422 S-palmitoyl cysteine (By similarity).
 FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
 FT CONFLICT 46 46 K -> N (in Ref. 2 and 3; AAB60875).
 FT CONFLICT 59 59 I -> T (in Ref. 3; AAB60873).
 FT CONFLICT 115 115 K -> B (in Ref. 1).
 FT CONFLICT 165 165 G -> V (in Ref. 3; AAB60873 and 4).
 FT CONFLICT 200 200 M -> V (in Ref. 2 and 3).
 FT CONFLICT 227 227 F -> L (in Ref. 3; AAB60873).
 FT CONFLICT 271 271 K -> B (in Ref. 3; AAB60873).
 FT CONFLICT 281 281 N -> H (in Ref. 3; AAB60873).
 SQ SEQUENCE 458 AA; 51158 MW; 17257 CRC64;

Query Match 89.6%; Score 1818; DB 1; Length 458;
 Best Local Similarity 89.3%; Pred. No. 7, 6e-125;
 Matches 352; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLLVQLALPAATQKRVVIGKKGVETLTCTASQKKSIOFHWKNSNOIK 60
 DB 1 MNWGIFFRHLILVQLALPAATQKRVVIGKKGVETLTCTASQKKSIOFHWKNSNOIK 60
 QY ILAGNOSPLTKGSKINDRADRSRLMDQGNFPLIKLKIENSDPTTICEVEDQKEVQL 120
 DB 61 ILKQKQSFRTKSSKLRDRDRSKSLMDQCFSMIIKNIKIEDESYICEVEKKEVEVL 120
 QY 121 LVFGLTANSPDTHLQSQSLTLTLSPGSSPSYQCRSPRGKNIQGGKTLVSQLELQDSG 180
 DB 121 LVFGLTANSPDTHLQSQSLTLTLSPGSSPSYQCRSPRGKNIQGGKTLVSQLELQDSG 180
 QY 181 TWTCYLVQNKQKVEFKIDIVLAFQKASSTIVYKKEGEVESPFLAFVTEKLTGSGELMW 240
 DB 181 TWTCYLVQNKQKVEFKIDIVLAFQKASSTIVYKKEGEVESPFLAFVTEKLTGSGELMW 240
 QY 241 QAEKASSSSKSWITFDLKNKEVSVKRYTODPKLOMGKLLPLHLTLPOLFOYVGSGLTLA 300
 DB 241 QAEKASSSSKSWITFDLKNKEVSVKRYTODPKLOMGKLLPLHLTLPOLFOYVGSGLTLA 300
 QY 301 LEAKTKLHQEVNLVYVRAATOLQKNLTCEVWGPTSPFKMLSLKLENKEAKVSKREKPVV 360
 DB 301 LEAKTKLHQEVNLVYVRAATOLQKNLTCEVWGPTSPFKMLSLKLENKEAKVSKREKPVV 360
 QY 361 LNPBGMQCLSDSQVULLESNIKVLPTWSTPV 394
 DB 361 LNPBGMQCLSDSQVULLESNIKVLPTWSTPV 394

RESULT 8
 CD4_ERYPA STANDARD; PRT; 397 AA.
 AC 008339;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3) (Fragment).
 GN Name=CD4;
 OS Erythrocytus patas (Red guenon) (Cercopithecus patas).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Erythrocytus.
 OX NCBI_TaxID=9538;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=93049640; PubMed=1425921;
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
 RT "Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus.";
 RL Eur. J. Immunol. 22:2973-2981(1992).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -----
 CC The SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X73324; CAA51750.1; -.
 DR HSSP: P01730; IWO.
 DR GO:GO:0042101; C:T-cell receptor complex; ISS.
 DR GO:GO:0015026; F:coreceptor activity; ISS.
 DR GO:GO:0042289; F:MHC class II protein binding; ISS.
 DR GO:GO:0006955; P:immune response; ISS.
 DR GO:GO:0045086; P:positive regulation of interleukin-2 biosyn. . . ; ISS.
 DR GO:GO:0030217; P:T-cell differentiation; ISS.
 DR GO:GO:0045058; P:T-cell selection; ISS.
 DR GO:GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; ISS.
 DR InterPro: IPR008424; CD2.
 DR InterPro: IPR000973; CD4 TCAG.
 DR InterPro: IPR003596; IG-like.
 DR Pfam: PF000477; IG_2.
 DR PRINTS: PR00692; CD4TCANTIGEN.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PS00835; IG LIKE; 1.
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein; Palmitate; Repeat; T-cell; Transmembrane.
 FT NON_TER 1 1
 FT DOMAIN 1 369
 FT TRANSMEM 370 391
 FT POTENTIAL 392 397
 FT DOMAIN 1 98
 FT DOMAIN 99 176
 FT DOMAIN 177 290
 FT DOMAIN 291 347
 FT CAROXYD 269 269
 FT CAROXYD 298 298
 FT DISULFID 14 82
 FT DISULFID 128 157
 FT DISULFID 301 343
 FT LIPID 392 392
 FT LIPID 395 395
 FT LIPID 397 397
 FT NON_TER 397 397
 SQ SEQUENCE 397 AA; 44081 MW; 67887397A6B7EA4F CRC64;

Query Match 84.6%; Score 1717; DB 1; Length 397;
 Best Local Similarity 89.9%; Pred. No. 1, 6e-117;
 Matches 330; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY 28 VLIAGKGDVVELTCTASQKKSIOFHWKNSNOIKILGNQGSFLTKGSKINDRADRSRLW 87
 DB 1 VLIAGKGDVVELTCTASQKKSIOFHWKNSNOIKILGNQGSFLTKGSKINDRADRSRLW 87
 QY 88 DQGNFPLIKNLKIEBDDTYICEVEDQKEVQLVRLTANSTHLLQSQSLTLTLSP 147
 DB 61 DQGNFPLIKNLKIEBDDTYICEVEDQKEVQLVRLTANSTHLLQSQSLTLTLSP 147
 QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGNTCTVQNKQKVEFKIDIVLAFQK 207
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGNTCTVQNKQKVEFKIDIVLAFQK 207
 QY 208 SSIYKKEGQVESPFLAFVTEKLTGSGELMWQAEKASSSSKSWITFDLKNKEVSVKRY 267
 DB 181 SSIYKKEGQVESPFLAFVTEKLTGSGELMWQAEKASSSSKSWITFDLKNKEVSVKRY 267
 QY 268 QDPKLOMGKLLPLHLTLPOLFOYVGSGLTLLEAKTKLHQEVNLVYVRAATOLQKNLT 327
 DB 241 QDPKLOMGKLLPLHLTLPOLFOYVGSGLTLLEAKTKLHQEVNLVYVRAATOLQKNLT 327

QY 328 CEVWGPTSPKLM,SLKLENKEAKVSRKRPVWVLPNBEAGMOCCLSDSGQVLLSNIKVL 387
 DB 301 CEVWGPTSPKLTSLKLENKEATISKQAKAVWVLPNBEAGMOCCLSDSGQVLLSNIKVL 360

QY 388 PTWSTPV 394
 DB 361 PTWSTPV 367

RESULT 9
 CD4_CERTO STANDARD; PRT; 397 AA.

AC 008336;
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3) (Fragment).
 GN Name=CD4;
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
 OC Cercopithecinae; Cercopithecus.
 NC NCBL_TaxID=9531;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=93049640; PubMed=1425921;
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
 RT "Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus." (1992).
 RL Eur. J. Immunol. 22:2973-2981(1992).

CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.1sb-sib.ch/announce/ or send an email to license@sib-sib.ch).

DR EMBL: X73328; CA51754.1; -;
 DR EMBL: X73327; CA51753.1; -;
 DR HSBP: F01730; IWIO.
 DR GO: GO:0042101; C:T-cell receptor complex; ISS.
 DR GO: GO:0015026; F:coreceptor activity; ISS.
 DR GO: GO:0042289; F:MHC class II protein binding; ISS.
 DR GO: GO:006955; P:immune response; ISS.
 DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO: GO:0030217; P:T-cell differentiation; ISS.
 DR GO: GO:0045058; P:T-cell selection; ISS.
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro: IPR008424; CD2.
 DR InterPro: IPR000973; CD4_Tcag.
 DR InterPro: IPR007110; Ig_V.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_2.
 DR PRINTS: PR00692; CD4TCANTIGEN.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PS00835; IG_LIKE; 1.
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein; Palmitate; Repeat; T-cell; Transmembrane.
 FT NON_TER 1
 FT DOMAIN <1 369 Extracellular (Potential).
 FT TRANSMEM 370 391 Potential.

FT DOMAIN 392 >397 Cytoplasmic (Potential).
 FT DOMAIN <1 98 Ig-like V-type.
 FT DOMAIN 99 176 Ig-like C2-type 1.
 FT DOMAIN 177 230 Ig-like C2-type 2.
 FT DOMAIN 291 347 Ig-like C2-type 3.
 FT CARBOHYD 15 15 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 269 269 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 298 298 N-linked (GlcNAc. . .) (Potential).
 FT DISULFID 14 82 By similarity.
 FT DISULFID 128 157 By similarity.
 FT DISULFID 301 343 By similarity.
 FT LIPID 392 392 S-palmitoyl cysteine (By similarity).
 FT LIPID 395 395 S-palmitoyl cysteine (By similarity).
 FT VARIANT 20 20 Missing.
 FT VARIANT 43 43 T -> I.
 FT VARIANT 86 86 N -> D.
 FT VARIANT 96 96 F -> L.
 FT VARIANT 173 173 V -> M.
 FT VARIANT 316 316 R -> K.
 FT NON_TER 397 397

SO SEQUENCE 397 AA; 43926 MW; 8660B636D2DB38A7 CRC64;

Query Match 84.1%; Score 1707; DB 1; Length 397;
 Best Local Similarity 89.6%; Pred. No. 8, 6e-117;
 Matches 329; Conservative 18; Mismatches 20; Indels 0; Gaps 0;

QY 28 VYLKGGDVELCTPSQKKSIOFHKNKNOIKILGNQSPFLTKPSKINDRADSRSLW 87
 DB 1 VYLKGGDVELACNQSQKKSIOFHKNKSKQIKILGNQSPFLTKSSKLSRADSRSLW 60

QY 88 DQGNFPLIKNIKIESDPTIYCEVEDQKEVOLVGLTANSPTHLLOQOSLTLTLESPP 147
 DB 61 DQGNFPLIKNIKIESDPTIYCEVEDQKEVOLVGLTANSPTHLLOQOSLTLTLESPP 120

QY 148 GSSPSVQCSPPKINIQGKTLVSQLEQDSGTWTCTYLQNKVKEFKIDIVIAFOKA 207
 DB 121 GSSPSVQCSPPKINIQGKTLVSQLEQDSGTWTCTYLQNKVKEFKIDIVIAFOKA 180

QY 208 SSTVYKKEGVOYEFSPPLAFYBKLTGSGELMWQARASSSWITPDLKKEVSKRYT 267
 DB 181 SSTVYKKEGVOYEFSPPLAFYBKLTGSGELMWQARASSSWITPDLKKEVSKRYT 240

QY 268 QPKLQMGKKLPLHLTLPOALPOYAGSNLTLALAKTGKLEHGVNLVVMRATOLQKNLT 327
 DB 241 QPKLQMGKKLPLHLTLPOALPOYAGSNLTLALAKTGKLEHGVNLVVMRATOLQKNLT 300

QY 328 CEVWGPTSPKLM,SLKLENKEAKVSRKRPVWVLPNBEAGMOCCLSDSGQVLLSNIKVL 387
 DB 301 CEVWGPTSPKLTSLKLENKEATISKQAKAVWVLPNBEAGMOCCLSDSGQVLLSNIKVL 360

QY 388 PTWSTPV 394
 DB 361 PTWSTPV 367

RESULT 10
 ID 009261 PRELIMINARY; PRT; 397 AA.
 AC 009261;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CD4 (Fragment).
 GN Name=CD4;
 OS Cercopithecus sabaues.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
 OC Cercopithecinae; Cercopithecus.
 NC NCBL_TaxID=60711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98017879; PubMed=9379478;
 RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
 RT "Relation between phylogeny of African green monkey CD4 genes and
 their respective simian immunodeficiency virus genes.";
 RL J. Med. Primatol. 26:120-128(1997).
 DR EMBL: AF001224; AAB60872.1; -.
 DR HSSP: P01730; 1WIO.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0007155; P:cell adhesion; IEA.
 DR GO: GO:0006955; P:immune response; IEA.
 DR InterPro: IPR008424; CD2.
 DR InterPro: IPR000973; CD4_TcAg.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig v.
 DR Pfam: PF05790; C2-set; 2.
 DR Pfam: PF00047; Ig; 1.
 DR PRINTS: PR00692; CD4TCANTIGEN.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PS50835; IG_Like; 1.
 FT NON_TER 1 1
 FT SEQUENCE 397 AA; 43945 MW; 80C963B92A868CD3 CRC64;
 SQ

Query Match 84.0%; Score 1705; DB 2; Length 397;
 Best Local Similarity 89.9%; Pred. No. 1.2e-116;
 Matches 330; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY 28 VVLGKGGDTVELCTASQKSIQPHWKNNSQIKILGNQGSFLTKGSPSKLNDRAISRSLW 87
 DB 1 VVLGKGGDTVELCTASQKSIQPHWKNNSQIKILGNQGSFLTKGSKLDRIDSRKSLW 60
 QY 88 DQGNFPLIINKLKIEDSDTYICEVEDQKEVQLVPLGTANSPTHLLOQGSLLTLESPP 147
 DB 61 DQGFSSMIINKLKIEDSEITYICEVENKEVEVLVPLGTANSPTHLLOQGSLLTLESPP 120
 QY 148 GSSPSVOCBPRGNIOGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIVVLAFOKA 207
 DB 121 GSSPSVOCRSPRGNIOGKTLSPQLERQDSGTWTCTVSDQNMVEFKIDIVVLAFOKA 180
 QY 208 SSIYKKEGEQVEFSPLAFTVEKLTGSGELMWQAEPASSSKSWITTDLNKKEVSVKVT 267
 DB 181 SSTYKKEGEQVEFSPLAFTVEKLTGSGELMWQAEPASSSKSWITTDLNKKEVSVKVT 240
 QY 268 QDPKIQMGKKLPILHTLPQALPOVAGSNLTALAEAKTGKHQEVNLVWMRATQLOKNT 327
 DB 241 QDPKIQMGKKLPILHTLPQALPOVAGSNLTALAEAKTGKHQEVNLVWMRATQLOKNT 300
 QY 328 CEVWGPTSPKMLSLKLENKAKYSKREKPVVNLNPEAGMOCCLSDSGVLLSNIKVL 387
 DB 301 CEVWGPTSPKMLSLKLENKAKATVSKQAKAVVNLNPEAGMOCCLSDSGVLLSNIKVL 360
 QY 388 PTWSTPV 394
 DB 361 PTWPTPV 367

RESULT 11
 ID 009260 PRELIMINARY; PRT; 397 AA.
 AC 009260;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE CD4 (Fragment).
 GN Name=CD4;
 OS Cercopithecus sabaues.
 OC Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Cercopithecus.
 OX NCBI_TaxID=60711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98017879; PubMed=9379478;
 RA Pomegaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
 RT "Relation between phylogeny of African green monkey CD4 genes and
 their respective simian immunodeficiency virus genes.";
 RL J. Med. Primatol. 26:120-128(1997).
 DR EMBL: AF001224; AAB60872.1; -.
 DR HSSP: P01730; 1WIO.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0007155; P:cell adhesion; IEA.
 DR GO: GO:0006955; P:immune response; IEA.
 DR InterPro: IPR008424; CD2.
 DR InterPro: IPR000973; CD4_TcAg.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig v.
 DR Pfam: PF05790; C2-set; 2.
 DR Pfam: PF00047; Ig; 1.
 DR PRINTS: PR00692; CD4TCANTIGEN.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE: PS50835; IG_Like; 1.
 FT NON_TER 1 1
 FT SEQUENCE 397 AA; 43882 MW; 478BB27E992EE89 CRC64;
 SQ

Query Match 83.8%; Score 1701; DB 2; Length 397;
 Best Local Similarity 89.6%; Pred. No. 2.4e-116;
 Matches 329; Conservative 16; Mismatches 22; Indels 0; Gaps 0;

QY 28 VVLGKGGDTVELCTASQKSIQPHWKNNSQIKILGNQGSFLTKGSPSKLNDRAISRSLW 87
 DB 1 VVLGKGGDTVELCTASQKSIQPHWKNNSQIKILGNQGSFLTKGSKLDRIDSRKSLW 60
 QY 88 DQGNFPLIINKLKIEDSDTYICEVEDQKEVQLVPLGTANSPTHLLOQGSLLTLESPP 147
 DB 61 DQGFSSMIINKLKIEDSEITYICEVENKEVEVLVPLGTANSPTHLLOQGSLLTLESPP 120
 QY 148 GSSPSVOCBPRGNIOGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIVVLAFOKA 207
 DB 121 GSSPSVOCRSPRGNIOGKTLSPQLERQDSGTWTCTVSDQNMVEFKIDIVVLAFOKA 180
 QY 208 SSIYKKEGEQVEFSPLAFTVEKLTGSGELMWQAEPASSSKSWITTDLNKKEVSVKVT 267
 DB 181 SSTYKKEGEQVEFSPLAFTVEKLTGSGELMWQAEPASSSKSWITTDLNKKEVSVKVT 240
 QY 268 QDPKIQMGKKLPILHTLPQALPOVAGSNLTALAEAKTGKHQEVNLVWMRATQLOKNT 327
 DB 241 QDPKIQMGKKLPILHTLPQALPOVAGSNLTALAEAKTGKHQEVNLVWMRATQLOKNT 300
 QY 328 CEVWGPTSPKMLSLKLENKAKYSKREKPVVNLNPEAGMOCCLSDSGVLLSNIKVL 387
 DB 301 CEVWGPTSPKMLSLKLENKAKATVSKQAKAVVNLNPEAGMOCCLSDSGVLLSNIKVL 360
 QY 388 PTWSTPV 394
 DB 361 PTWPTPV 367

RESULT 12
 ID 095NE9 PRELIMINARY; PRT; 397 AA.
 AC 095NE9;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE CD4 (Fragment).
 GN Name=CD4;
 OS Cercopithecus pygerythrus (Vervet monkey).
 OC Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Cercopithecus.
 OX NCBI_TaxID=60710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98017879; PubMed=9379478;
 RA Pomegaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001227; AAB60874.1; -.
DR HSSP: P01730; 1WIO.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0007155; P: cell adhesion; IEA.
DR GO: GO:0006955; P: immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcRg.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-sect. 2.
DR Pfam: PF00047; Ig. 1.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV. 1.
DR PROSITE: PS50835; IG_LIKE. 1.
FT NON TER 1
FT 397
SQ SEQUENCE 397 AA; 43946 MW; 21C3E3082ABPBC0 CRC64;

Query Match 83.8%; Score 1700; DB 2; Length 397;
Best Local Similarity 89.6%; Pred. No. 2.8e-116;
Matches 329; Conservative 15; Mismatches 23; Indels 0; Gaps 0;

QY 28 VVLGKKGDVVELTCTASOKKSIQFHMKNSNOIKIIGNGSFLTKGSKLNDPADSRSLW 87
DB 1 VVLGKKGDVVELTCTASOKKSIQFHMKNSNOIKIIGNGSFLTKGSKLNDPADSRSLW 60
QY 88 DQGNFPLIKNIKIEDSDTYICEVEDQKEEVOLVFGLTANSPTHLLOGQSILTLTLESP 147
DB 61 DQGCFSMIITKNKIEDSEYICEVENKEKEVELVFGLTANSPTHLLOGQSILTLTLESP 120
QY 148 GSSPSVQCKSPKKNIOGGKITLSVSOLELDSDGTCTCYLQNKVKEFKIDIVIAFOKA 207
DB 121 GSSPSVKCKSPKKNIOGGKITLSVPOLEKSDGTCTCWSQDNVVEFKIDIVIAFOKA 180
QY 208 SSIYVKKEGEQVEFPLAFTVEKLTGSGELMWQAEARSSSKSWITFDLKNKEVSKRYT 267
DB 181 SSIYVKKEGEQVEFPLAFTVEKLTGSGELMWQAEARSSSKSWITFDLKNKEVSKRYT 240
QY 268 QDPKIQMGKPLNLTLPQALPOYAGSGNLTLALBAKTKLHQEVNLVVMRAITOLQKNLT 327
DB 241 QDPKIQMGKPLNLTLPQALPOYAGSGNLTLALBAKTKLHQEVNLVVMRAITOLQKNLT 300
QY 328 CEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDSGVLLSNIKYL 387
DB 301 CEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDSGVLLSNIKYL 360
QY 388 PTWSTPV 394
DB 361 PTWSTPV 367

RESULT 13

ID 009259 PRELIMINARY; PRT; 397 AA.
AC 009259;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD4 (Fragment).
GN Name=CD4;
OS Cercopithecus babuinus.
OC Cercopithecidae; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_Taxid=60711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001223; AAB60870.1; -.
DR HSSP: P01730; 1WIO.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0007155; P: cell adhesion; IEA.
DR GO: GO:0006955; P: immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcRg.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-sect. 2.
DR Pfam: PF00047; Ig. 1.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV. 1.
DR PROSITE: PS50835; IG_LIKE. 1.
FT NON TER 1
FT 397
SQ SEQUENCE 397 AA; 43881 MW; 7CE39AD0F8506C81 CRC64;

Query Match 83.6%; Score 1696; DB 2; Length 397;
Best Local Similarity 89.4%; Pred. No. 5.5e-116;
Matches 328; Conservative 17; Mismatches 22; Indels 0; Gaps 0;

QY 28 VVLGKKGDVVELTCTASOKKSIQFHMKNSNOIKIIGNGSFLTKGSKLNDPADSRSLW 87
DB 1 VVLGKKGDVVELTCTASOKKSIQFHMKNSNOIKIIGNGSFLTKGSKLNDPADSRSLW 60
QY 88 DQGNFPLIKNIKIEDSDTYICEVEDQKEEVOLVFGLTANSPTHLLOGQSILTLTLESP 147
DB 61 DQGCFSMIITKNKIEDSEYICEVENKEKEVELVFGLTANSPTHLLOGQSILTLTLESP 120
QY 148 GSSPSVQCKSPKKNIOGGKITLSVSOLELDSDGTCTCYLQNKVKEFKIDIVIAFOKA 207
DB 121 GSSPSVKCKSPKKNIOGGKITLSVPOLEKSDGTCTCWSQDNVVEFKIDIVIAFOKA 180
QY 208 SSIYVKKEGEQVEFPLAFTVEKLTGSGELMWQAEARSSSKSWITFDLKNKEVSKRYT 267
DB 181 SSIYVKKEGEQVEFPLAFTVEKLTGSGELMWQAEARSSSKSWITFDLKNKEVSKRYT 240
QY 268 QDPKIQMGKPLNLTLPQALPOYAGSGNLTLALBAKTKLHQEVNLVVMRAITOLQKNLT 327
DB 241 QDPKIQMGKPLNLTLPQALPOYAGSGNLTLALBAKTKLHQEVNLVVMRAITOLQKNLT 300
QY 328 CEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDSGVLLSNIKYL 387
DB 301 CEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDSGVLLSNIKYL 360
QY 388 PTWSTPV 394
DB 361 PTWSTPV 367

RESULT 14

ID 009263 PRELIMINARY; PRT; 397 AA.
AC 009263;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD4 (Fragment).
GN Name=CD4;
OS Cercopithecus tantalus.
OC Cercopithecidae; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_Taxid=60712;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.:
RT "Relation between phylogeny of African green monkey CD4 genes and
their respective simian immunodeficiency virus genes.",
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL; AF001221; AAB60869.1; -.
DR HSSP; P01730; 1WIO.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR00973; CD4_TcRg.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF05790; C2-set; 2.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 397
SQ SEQUENCE 397 AA; 43994 MW; A3CD031535A51524 CRC64;

Query Match 83.6%; Score 1696; DB 2; Length 397;
Best Local Similarity 89.4%; Pred. No. 5.5e-116;
Matches 328; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

```
QY 28 VVLGGKGDVVELTCTASQKKSIOFHMKNNOIKILGNQGSFLTQGPSKLNDRADSRSLW 87
DB 1 VVLGGKGDVVELTCTASQKKSIOFHMKNNOIKILGNQGSFLTQGPSKLNDRADSRSLW 60
QY 88 DQGNFPLIIKNLIKIEDSDTYICEVEDQKEBVOLVFGLTANSPTHLLOQSITLTLSP 147
DB 61 DQGCFSMIIKNIKIEDSETYICEVENKKEVELLVFGLTANSPTHLLOQSITLTLSP 120
QY 148 GSSPSVQCSPPKKNIOGKKTLSVSLLELODSGTCTVLOKRYEFDIVLAFOKA 207
DB 121 GSSPSVKCSPPKKNIOGKKTLSVSLLELODSGTCTVLOKRYEFDIVLAFOKA 180
QY 208 SSIIVYKKEGEQVEFSFPLAFTVEKLTGSGELMWQARASSSKSMITFDLKNKEVSKRYT 267
DB 181 SSIIVYKKEGEQVEFSFPLAFTVEKLTGSGELMWQARASSSKSMITFDLKNKEVSKRYT 240
QY 268 QDPKLOMGKKLPILHTLTPQALPOYAGSGNLTALFAKTGKLHOEVNLVVMRATOLQKNLT 327
DB 241 QDPKLOMGKKLPILHTLTPQALPOYAGSGNLTALFAKTGKLHOEVNLVVMRATOLQKNLT 300
QY 328 CEVWGTSPEKMLSLKLENKAKVSKREKPVWVLNPEAGMOCCLSDSGQVLLSENIKYL 387
DB 301 CEVWGTSPEKMLSLKLENKAKVSKREKPVWVLNPEAGMOCCLSDSGQVLLSENIKYL 360
QY 388 PTWSTPV 394
DB 361 PTWSTPV 367
```

RESULT 15
009262 PRELIMINARY; PRT; 397 AA.

AC 009262;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CD4 (Fragment).
GN Name=CD4;
OS Cercopithecus tanzianus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
OX NCBI_Taxid=60712;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.:
RT "Relation between phylogeny of African green monkey CD4 genes and
their respective simian immunodeficiency virus genes.",
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL; AF001221; AAB60869.1; -.
DR HSSP; P01730; 1WIO.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR00973; CD4_TcRg.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF05790; C2-set; 2.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 397
SQ SEQUENCE 397 AA; 43994 MW; CF7F2F5D82335B0D CRC64;

Query Match 83.2%; Score 1689; DB 2; Length 397;
Best Local Similarity 89.4%; Pred. No. 1.8e-115;
Matches 328; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

```
QY 28 VVLGGKGDVVELTCTASQKKSIOFHMKNNOIKILGNQGSFLTQGPSKLNDRADSRSLW 87
DB 1 VVLGGKGDVVELTCTASQKKSIOFHMKNNOIKILGNQGSFLTQGPSKLNDRADSRSLW 60
QY 88 DQGNFPLIIKNLIKIEDSDTYICEVEDQKEBVOLVFGLTANSPTHLLOQSITLTLSP 147
DB 61 DQGCFSMIIKNIKIEDSETYICEVENKKEVELLVFGLTANSPTHLLOQSITLTLSP 120
QY 148 GSSPSVQCSPPKKNIOGKKTLSVSLLELODSGTCTVLOKRYEFDIVLAFOKA 207
DB 121 GSSPSVKCSPPKKNIOGKKTLSVSLLELODSGTCTVLOKRYEFDIVLAFOKA 180
QY 208 SSIIVYKKEGEQVEFSFPLAFTVEKLTGSGELMWQARASSSKSMITFDLKNKEVSKRYT 267
DB 181 SSIIVYKKEGEQVEFSFPLAFTVEKLTGSGELMWQARASSSKSMITFDLKNKEVSKRYT 240
QY 268 QDPKLOMGKKLPILHTLTPQALPOYAGSGNLTALFAKTGKLHOEVNLVVMRATOLQKNLT 327
DB 241 QDPKLOMGKKLPILHTLTPQALPOYAGSGNLTALFAKTGKLHOEVNLVVMRATOLQKNLT 300
QY 328 CEVWGTSPEKMLSLKLENKAKVSKREKPVWVLNPEAGMOCCLSDSGQVLLSENIKYL 387
DB 301 CEVWGTSPEKMLSLKLENKAKVSKREKPVWVLNPEAGMOCCLSDSGQVLLSENIKYL 360
QY 388 PTWSTPV 394
DB 361 PTWSTPV 367
```

Search completed: March 7, 2005, 07:19:30
Job time : 151.542 secs